Procurement and Contracting Services

Request for Proposals for Master Data Management

ADDENDUM #1

Please mark all proposal submission Envelopes with the following information

Sealed RFP # L162103
Due on 7/19/21 no later than 2:00 PM, MST
The purpose of this addendum is to answer the submitted questions. Responses in red.

1) **Architecture and Integration.** Can you please share a high-level architecture of the MDM solution if there is any? Specifically, are you aiming for Consolidation style MDM initially? Could you please clarify the systems consuming data from MDM? Should we assume batch and/or API integrations with data sources/data consumers?
   a) The goal will be consolidation style. Initially salesforce will be the primary consumer of data from MDM. API (web services) are the preferred integration method, but batch would be acceptable.

2) What is your technical environment now (Operating Systems, COTS System, Databases, Datawarehouse, DataMarts )
   a) Our technical environment is extremely varied. Generally speaking, systems we still host ourselves are in AWS. That would include our SIS, Peoplesoft, and our financial system, Kuali. All of our ERP systems are Oracle based. Our primary data warehouse is OBIEE+ using Datastage for ETL. Other systems are SaaS offerings, such as Salesforce, ServiceNow, o365, and more.

3) Other than Salesforce, what other ERP, COTS application do the University currently using
   a) Kuali Financials, Kuali Research, Peoplesoft Student, and Peoplesoft HCM are our primary ERP systems.

4) Does the University have a preferred Cloud Service Provider for hosting the SaaS solution, or you looking for a PaaS solution (both Cloud Platform and Software as a bundle)
   a) We prefer AWS, but do have some Azure resources as well.

5) What ETL/middleware tools are currently used?
   a) We use Datastage for our data warehouse. Much of the integration for Salesforce is done through Streamsets. There is some use of AWS Glue, and a smattering of AWS Lambda functions as well.

6) Integration requirements with each source: batch and/or real-time
   a) How many systems are sourcing data in batch mode? Details of the frequency for each system.
      i) It will depend on the flexibility of the solution. Most likely all systems will be batch mode.
   b) How many systems are sourcing data in real time?
      i) It will depend on the flexibility of the solution. Most likely only user created records throughout the day would be sourced in real time.
   c) What is the volume of data for batch integration?
      i) Approximately 500,000 records.
   d) What is the volume of data for real time integration?
      i) Thousands of inserts/day at most.

7) What is your preferred cloud hosting service provider (Azure or AWS)?
   a) AWS
8) Can we utilize a hybrid resource model (onsite and offshore resources) to complete the project? 
   a) Yes

9) Can we use an onsite-offshore model to deliver the services utilizing our resources based in India? 
   a) Yes

10) Does the University of Arizona prefer any specific product stacks like Microsoft, Oracle, Informatica etc? 
    a) Our focus will be on the MDM capabilities, with no specific allegiance to existing product stacks.

11) How are the inbound and outbound data sources connecting to this solution? 
    a) It will depend on what connectivity options the solution offers.

12) Delivery model. Do you prefer Vendor to deliver the initial onboarding (stand-up environment, training) or Phase 1 of the solution (minimum viable product or other), or the full solution (onboarding/phase1/rollout)? 
    a) Our preference will ultimately be determined by price of option. If you have those three delivery models, we would very much appreciate having each priced out for us.

13) Data model. Do you already have a draft of business/logical data model that can be shared, or will data model be created within the project? Could you please roughly estimate how many entities and attributes will be covered? 
    a) There is no universal logical data model that applies here. Some of it is built out, but some of it will need to be built as part of this project. 
    b) Roughly 1.5 million entities at the outset, and roughly 50 attributes.

14) Legacy system migration. Is migration from a legacy system required? 
    a) No.

15) What are business drivers for the Master Data Management Project for the University 
    a) MDM will be the solution to the problem of bringing data from disparate sources around the University together to be cleaned and. It will be the centerpiece of the process to identify the “golden record,” the best available contact record. We seek to solve the following problems: 
       i) Inconsistency 
       ii) Inaccuracy 
       iii) Timeliness 
       iv) Non-standard 
       v) Duplicity 
       vi) Unmanaged (no rules) 
       vii) Hard to find constituents 
       viii) Improper grouping of constituents
16) How many business domains University like to be part of the MDM project
   a) If a business domain is defined as group of like data, then 2, Accounts/Businesses and
      People/contacts. We will build the MDM with an eye toward potentially scaling to include
      more domains in the future.

17) What are critical use cases for Maser Data Management specifically for the University
   a) We seek to solve the following problems:
      i) Inconsistency
      ii) Inaccuracy
      iii) Timeliness
      iv) Non-standard
      v) Duplicity
      vi) Unmanaged (no rules)
      vii) Hard to find constituents
      viii) Improper grouping of constituents
      ix) Poor system usability
      x) Inability to leverage data & find insights
      xi) Lack of data ownership & stewardship

18) Is this project has approved budget, and what is the range of the including implementation effort
   a) no budget information is to be provided

19) How much expected volume of Golden record for the Master Data Management (Approximate will
    work)
   a) 1.3 million at the outset to increase

20) What is the ideal implementation time the University is expecting to implement the MDM solution
   a) Ideally by the end of 2021.

21) Are additional data management initiatives such as data governance, metadata, data catalog, data
    quality, and data strategy progress? If yes, can you list product information
   a) No.

22) Has University seen any vendor product demos before the RFP? If yes, can you share the names
   a) Yes, Ataccama and Boomi

23) Will the University entertain teaming proposals?
   a) Yes.

24) Can you able to share in information of the participants attending pre-bid conference
   a) We did not take attendance at the pre-bid conference or record the session, so we are
      unable to share the participants.

25) The RFP identifies data mastering objectives for “People”. How are these categorized? (please note
    that this list is not exhaustive)
   a) Students: Active, Inactive, Former, Applicants, prospects
b) Faculty and Administrative: staff, faculty, dcc, student-workers

c) Other: Donors, Alumnae, ticket holders, former-members

26) The RFP identifies data mastering objectives for “Accounts” How are these categorized? Examples include:
   a) Vendors, Universities, Departments, Contractors, Sponsors, ticket buyers, recruiters, co-
    awardees

27) How many data sources are there? The RFP indicates disparate.
   a) There are 4 at the outset with at least 2 more on the horizon with the expectation to scale.

28) Are they on prem or cloud based?
   a) Cloud

29) What other applications may be relevant to a data mastering solution, beyond Salesforce?
   a) Blackbaud, PeopleSoft, Slate, Paciolan initially

30) Is UA open to a phased approach to project implementation?
   a) Yes.

31) Data Volumes:
   a) What are the total number to constituents that will be mastered in MDM?
      i) Initially there will be about 1 million will intention to scale.
   b) How many years of constituent data?
      i) About 20?
   c) Is Alumni included within this total?
      i) no.
   d) What are the total number of foreign students?
      i) about 2,800
   e) How many new constituents are added per year?
      i) estimated to be about 20,000
   f) Are applicants in scope for mastering?
      i) Yes.
   g) What is the volume of addresses per constituent to be mastered?
      i) We anticipate 3.
   h) What is the volume of phone numbers per constituent to be mastered?
      i) We anticipate 3.
   i) What is the volume of email addresses per constituent to be mastered?
      i) We anticipate up to 5.

32) Data Quality
   a) What is the current level of data quality of the data in sources (including Snowflake)?
      i) extremely variable by data source.
   b) What is the level of quality of addresses, emails, birth dates and phones within the sources?
      i) extremely variable by data source.
   c) What is the current level of quality for Metadata (ie countries, states, address, phone, email
    types, ethnicity, etc)?
      i) Our metadata is quite clean and well maintained for the initial systems in question,
   d) Are standards or tools in place to enforce Global standards for address, email, phone?
      i) No.
e) Are third party solutions used to maintain accurate Alumni data?
   i) Not that I am aware of, but possibly.

33) Data Governance
   a) Do you currently have a data governance structure in place, including data stewardship, formal data governance organization, escalation process, operating model, policies and procedures, etc.?
      i) We have an emerging data governance structure that varies across the university.
   b) What are the data governance policies/practices that are in place? Please provide the details. What is the volume of policies?
      i) Not applicable.

34) Goals & Scope
   a) What are the overall goals for implementing a MDM Solution?
      i) We seek to solve the following problems:
          (1) Inconsistency
          (2) Inaccuracy
          (3) Timeliness
          (4) Non-standard
          (5) Duplicity
          (6) Unmanaged (no rules)
          (7) Hard to find constituents
          (8) Improper grouping of constituents
          (9) Poor system usability
          (10) Inability to leverage data & find insights
          (11) Lack of data ownership & stewardship
   b) What is the scope of an implementation (by phase)?
      i) Phase 1 should focus on the following tasks:
          (1) Build MDM environment including all Extraction Transfer Load (ETL) packages
          (2) Define and design data warehouse structure utilizing data from each data source
          (3) Configure underlying data management and master data management systems to incorporate initial data sources
          (4) Implement data cleansing process
          (5) Configure/develop data integration services
          (6) Test integrations for accuracy with UA
   c) Would this solution be replacing an existing solution?
      i) No.

35) Sources
   a) Please identify the source types: Ie Salesforce, WorkDay, etc
      i) Salesforce, Blackbaud, PeopleSoft, Slate, Paciolan initially

36) Source Systems
   The RFP lists 5 source systems that are in scope for implementation () + several .csv files
a) What is the estimated number of tables per source system relevant for matching/deduplication? For each of the source systems can you give an estimate of the size of the relevant tables?
   i) 2 tables on average per source system
   ii) Paciolan – don't know
   iii) Salesforce – ~550k records with roughly 50 fields
   iv) Blackbaud – don't know
   v) Slate - ~1.3 million records with roughly 25 fields
   vi) Peoplesoft - ~350k records with about 35 fields
b) If possible, could you indicate the way we would be able to connect to each source system (i.e., through connecting to db, API, Webhook, etc.)?
   i) Connectivity options vary significantly, and we try to be as flexible as possible. Direct database connections will not be an option. If the tool provides native support for various tools and can connect directly via API that is an option. Otherwise we will utilize our integration layer, Streamsets, as necessary.
c) What type of integration is expected for OBIEE (should the MDM solution provide the data available for reporting purposes, or is there a DWH where the solution should provide data to)?
   i) At this point no integration with our local reporting is currently in scope. We would expect some amount of reporting from the selected tool itself.
d) How many additional .csv files should be factored into the design?
   i) 2 at the outset with room to scale.

37) Data Governance:
   a) During the pre-RFP call you expressed an interest in data governance.
      i) yes.
   b) Do you expect the MDM platform to play a role in / facilitate the data governance organization or do you see the two as distinct?
      i) Both and. We would like the MDM platform to allow for workflows and permissions based on data governance roles
   c) Is there an existing data governance organization in place (with roles and responsibilities)? If so, what is currently implemented?
      i) We have an emerging data governance structure that varies across the university.
   d) Would you like the MDM solution to provide DQ dashboards to enable insight into the quality of Constituent data?
      i) Yes please.

38) Number of records
   a) During the pre-RFP call we discussed the number of records (expected around 1.5 million). At the time of the call you were not sure about this number.
      i) Would you at least be able to provide a rough estimate (order of magnitude) for the number of consolidated records that can be expected?
         (1) Roughly 1.3 million.

39) Section 5.2
   a) In section 5.2 a table is provided with a comprehensive list of technical requirements. Would you be able to provide an editable version of this table?
i) Yes.
b) Having an editable version would save us a significant amount of time when indicating whether our solution has that in base functionality, configurable, etc. – above

40) Do you expect the solution to be a COTS system that satisfies entire requirements, or shall it be customized, or custom developed as per the University’s business and technical needs?
a) It can be either depending on what best fits our needs and price point.

41) Can you provide us the list of data sources?
a) Blackbaud, PeopleSoft, Slate, Paciolan initially

42) Do all of your data sources have an active data dictionary?
a) No.

43) Do you have documented data quality rules/standards?
a) No.

44) What are the master data entities that you would like to have?
a) I don’t know what ‘master data entities’ refers to exactly. We would like to have composite records for each person/contact and business/account.

45) Can you describe any hierarchies within the domains or related systems? Are they complex? Can you share examples?
a) At the moment, there are no person hierarchies. We would like to be able to create household hierarchies though. There are typical parent-subsidary hierarchies for businesses/accounts. They are not particularly complex.

46) How are you currently manage metadata? Do you have any existing tools?
a) We are not systematically managing metadata, and we have no existing tools.

47) Do you expect to use any external business intelligence tools for visualization?
a) OBIEE, Tableau

48) Page 30: “Automatically trigger email notification on data onboarding” Is this expected whenever new data is pulled from source systems? Can you please mention the frequency?
a) No, but if a new table or large chunk of new data is pulled, we would like that to trigger the email. The exact specifications are malleable.

49) Considering mid of August as the contract start date, what is the anticipated go-live date?
a) Ideally by the end of 2021.

50) Can you provide us the ballpark figure for this project that we need to consider?
a) no budget information is to be provided

51) Can the work be performed remotely for this project? If so, is it mandatory for the resource(s) to be residing in the US?
a) Yes
b) No
52) Is there a budget allocated for this project? Is there an upper limit we should be aware of?
   a) no budget information is to be provided

53) Is there any format that we need to follow while preparing the proposal?
   a) No. Please use whatever format conveys your proposed solution in the most reader friendly way.

54) Page 24: “4.28 Intellectual Property. It is understood and agreed that ownership of intellectual property developed as a result of fulfilling the requirements of this Request for Proposals belongs solely and exclusively to the Arizona Board of Regents on behalf of the University of Arizona.”
   a) Do you expect to own the IP rights for the overall MDM solution?
      i) We do not. Anything software or system that is developed specifically for UA use would be considered UA IP, but a system/platform/product/service already in existence applied to UA would not be considered our IP.

55) Section 3.8.2 states, "Under this RFP, the University shall hold that vendors' responses to this RFP shall remain in effect for a period of ninety (90) days following the Due date, in order to allow time for evaluation, approval, and award of the contract." In referring to pricing, Section 5.5. states, "Any quotes must be valid for six (6) months." Can the University please clarify the period that vendors must certify that their responses are valid?
   a) The RFP response is separate from the quote, the response is good for 90 days and any resulting quote is good from 6 months.

End of addendum, all else remains the same.